

Store-Forward and its implications for Proportional Scheduling

N.S. Walton¹

Abstract—The Proportional Scheduler was recently proposed as a scheduling algorithm for multi-hop switch networks. For these networks, the BackPressure scheduler is the classical benchmark. For networks with fixed routing, the Proportional Scheduler is maximum stable, myopic and, furthermore, will alleviate certain scaling issues found in BackPressure for large networks. Nonetheless, the equilibrium and delay properties of the Proportional Scheduler has not been fully characterized.

In this article, we postulate on the equilibrium behaviour of the Proportional Scheduler through the analysis of an analogous rule called the Store-Forward allocation. It has been shown that Store-Forward has asymptotically allocates according to the Proportional Scheduler. Further, for Store-Forward networks, numerous equilibrium quantities are explicitly calculable. For FIFO networks under Store-Forward, we calculate the policies stationary distribution and end-to-end route delay. We discuss network topologies when the stationary distribution is product-form, a phenomenon which we call *product form resource pooling*. We extend this product form notion to independent set scheduling on perfect graphs, where we show that non-neighbouring queues are statistically independent. Finally, we analyse the large deviations behaviour of the equilibrium distribution of Store-Forward networks in order to construct Lyapunov functions for FIFO switch networks.

I. INTRODUCTION

Switch networks model numerous communication networks where service at a queue may inhibit the operation of other queues. Common examples include wireless ad-hoc networks [20] and input queued switches [14]. Further the paradigm is provides useful in insights in other contexts: bandwidth sharing, call centers, road traffic, data centers, manufacturing systems. At each time a scheduler makes a scheduling decision and, since load often varies, scheduling is often based on current or recent queue size information, rather than on explicit estimation of the network's long-run load. An important first order property of a “good” scheduling policy is to be *maximum stable* which states that, for the largest possible set of arrival rates, queues remain bounded (on average) and thus have an equilibrium length. Once this first order stability condition is established then secondary properties such as equilibrium queue length and delay can be investigated.

The celebrated BackPressure policies of Tassiulas and Ephremides [20] were the first class of scheduling policies that were proven to be maximum stable without explicit estimation of traffic load. The BackPressure policies are defined through a Lyapunov function argument, and provide

a robust, generic approach to stabilizing a queueing network. Nonetheless, characterizing the stationary behaviour of the resulting processes can be difficult, most bounds on stationary queue length are made through the policy's associated Lyapunov function [15], [17]. Further, delay scaling along lengthy routes can be shown to be suboptimal, see [4].

A scheduling policy called the Proportional Scheduler was recently proven to be maximum stable for multi-hop switch networks, for networks with fixed routing. The Proportional Scheduler has structural advantages in comparison to BackPressure. Packets can be aggregated at first-in first-out (FIFO) queues and, as a result, the scheduler does not maintain information about the routes taken by packets in order to make a scheduling decision. Since, for a communication network the number of routes supported by a network is, in general, orders of magnitude larger than the number of queues maintained by the network, this (in comparison to BackPressure) substantially reduces the structure required in memory to implement the policy and also leads to greater potential for decomposed implementation. Modifications of BackPressure have been developed in an effort to alleviate these issues [4], [6]. These and further consequences Proportional Scheduling are discussed in greater detail in the papers [24], [2].

The principle aim of this article is to propose queue length and delay estimates for the Proportional Scheduler, through formula that can be explicitly calculated for Store-Forward networks – a continuous time quasi-reversible queueing network which is known to asymptotically allocate resources as a Proportional Scheduler [23]. Due to the insensitivity property, the Store-Forward allocation was first analyzed by Bonald and Proutiere [1]. Indeed, the reason that the Proportional Scheduler mitigates some of the scaling phenomena found in BackPressure, is its close relationship between proportional fairness and reversible queueing systems (like the Store-Forward allocation). In reversible queueing networks, different packet types can be aggregated at each resource, and delays are explicit and depend on network load – rather than network location as is found in BackPressure.

We extend the single-class Markovian routing (or Jackson routing) considered by Bonald and Proutiere to a multi-class fixed routing (or Kelly routing). Through standard quasi-reversibility arguments we calculate the stationary distribution, queue size and delay for the Store-Forward Network.

Outline of Results

In Theorem 1, we calculate the stationary distribution of Store-Forward networks with fixed routes where queues

¹Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands. The work of Neil Walton is funded by the VENI research programme, which is financed by the Netherlands Organisation for Scientific Research (NWO). n.s.walton@uva.nl

operate under a FIFO discipline. We then calculate the stationary delay of these Store-Forward networks, Theorem 2, and thus postulate the stationary delay for switch queueing networks operating under the Proportional Scheduler.

Next, we develop on the product from resource pooling work found for bandwidth sharing networks by Kang, Kelly, Lee, Williams [7], [8]. For this work generalized Heavy traffic analysis and investigations into insensitivity are given by [26], [21]. First, it was argued in [18], [2] that independent switch components have independent stationary distributions, in heavy traffic. We extend these arguments to investigate Store-Forward on the interference graphs where the interference graph is perfect, and we show that the stationary queue length between non-adjacent queues are independent. See Theorem 3 and Corollary 2. The result is somewhat surprising given that independence is typically proven on product spaces.

Finally, we analyse the large deviations behaviour of a stationary Store-Forward Network. We discuss how the rate function found by this large deviations analysis can be used as a Lyapunov function for FIFO networks operating under the Proportional Scheduler.

II. QUEUEING NETWORK NOTATION

We let \mathcal{J} index a set of queues. A schedule σ is a vector in $\mathbb{Z}_+^{\mathcal{J}}$ and we let \mathcal{S} be the (finite) set of schedules. We assume that the set \mathcal{S} is monotone in the sense that if $\sigma \in \mathcal{S}$ and if $\tilde{\sigma} \in \mathbb{Z}_+^{\mathcal{J}}$ is $\tilde{\sigma} \leq \sigma$ component-wise then $\tilde{\sigma} \in \mathcal{S}$.

We let $\langle \mathcal{S} \rangle$ be the convex combination of points in \mathcal{S} . The set $\langle \mathcal{S} \rangle$ is a polytope and thus by our monotone assumption there exist a non-negative (full row rank) matrix $\mathbf{A} = (A_{lj} : l \in \mathcal{L}, j \in \mathcal{J})$ such that

$$\langle \mathcal{S} \rangle = \left\{ \mathbf{s} \in \mathbb{R}_+^{\mathcal{J}} : \sum_{j \in \mathcal{J}} A_{lj} s_j \leq 1, l \in \mathcal{L} \right\}.$$

We call each facet $l \in \mathcal{L}$ of the polytope $\langle \mathcal{S} \rangle$ a resource pool and, thus, \mathcal{L} is the set of resource pools. We use the notation that $j \in l$ if $A_{jl} > 0$ and $j \notin l$ if $A_{jl} = 0$.

A route is a finite ordered set of queues $r = (j_1^r, \dots, j_{k_r}^r)$. We will assume that the queues visited on a route r are distinct.¹ We let the finite set \mathcal{R} be the set of routes. We use the notation $j \in r$ if queue j is an element of route r .

We let $\mathbf{Q} = (Q_j : j \in \mathcal{J}) \in \mathbb{Z}_+^{\mathcal{J}}$ give the vector of queue sizes. We must also consider the order of jobs within a queue. We let $\Gamma_{jr} = (\Gamma_{jr}(1), \dots, \Gamma_{jr}(Q_j))$ counts the cumulative number of packets of route r at queue j , that is $\Gamma_{jr}(1) = 1$ if the first job in queue j is from route r , and equal zero otherwise. Similarly, $\Gamma_{jr}(2) - \Gamma_{jr}(1)$ indicates the presence of a route r job in the second position and so forth.

We let a_r be the arrival rate on route r . Thus the load on queue $j \in \mathcal{J}$ and the load on resource pool $l \in \mathcal{L}$ are respective given by

$$a_j := \sum_{r: j \in r} a_r, \quad a_l := \sum_{j \in l} A_{lj} a_j. \quad (1)$$

¹With a little extra notation, it is possible to visit a queue multiple times on a route.

III. THE STORE-FORWARD NETWORKS

We describe a continuous time Markov chain called the Store-Forward Network, see Section III-A. Here, queueing resources are allocated according to the Store-Forward allocation, rather than BackPressure or Proportional Fairness (cf. Section III-D). The desirable properties that are associated with Store-Forward – such as quasi-reversibility and product-form – are inherited from its close relationship with quasi-reversible queueing networks. Essentially the Store-Forward allocation is defined as the throughput of a certain Kelly network. This point is discussed below in Section III-B.

Further, properties found in Store-Forward pass over to the Proportional Scheduler. This is because the asymptotic behaviour of Store-Forward is to allocate resources in the same way as the Proportional Scheduler and, as we discussed in the introduction, the Proportional Scheduler has been shown to have a number of significant structural advantages in comparison to BackPressure. These points are discussed in more detail in Sections III-C and III-D.

We are interested in Store-Forward because of its implications for switch networks. Concrete implementations of Store-Forward in switch networks are possible, for instance, see [18] where optimal queue size bounds for Input-queued switches can be proven. However, we do not discuss direct implementation of Store-Forward in this article since we are interested in its consequences for the associated Proportional Scheduler. We discuss how results on Store-Forwards Networks can be phrased in the context of switch networks. Initial comments in this regard are made in Section III-E below, before developing this point further in subsequent sections.

A. Store-Forward Networks

Given a vector of queue sizes $\mathbf{Q} = (Q_j : j \in \mathcal{J})$ the Store-Forward allocation $\sigma^{SF}(\mathbf{Q})$ is defined by

$$\sigma_j^{SF}(\mathbf{Q}) = \frac{\Phi(\mathbf{Q} - \mathbf{e}_j)}{\Phi(\mathbf{Q})}, \quad (2)$$

for each $j \in \mathcal{J}$, where $\mathbf{e}_j \in \mathbb{Z}_+^{\mathcal{J}}$ is the j th unit vector. The positive function $\Phi(\mathbf{Q})$ has an explicit form; however, rather than give a formula for $\Phi(\mathbf{Q})$ which at this point would be uninformative, we delay its definition until Section III-B below.

The Store-Forward allocation was first introduced by Bonald and Proutiere [1] and has typically been studied as a model of bandwidth sharing. For this reason, packets at a queue leave after completing service and the discipline within a queue has typically been processor-sharing. However, in the context of switch networks, we wish packets to traverse between the queues on their route and for packets to be served in a first-in first-out (FIFO) manner. For this reason, we consider the following continuous time Markov chain, which we call a *FIFO routed Store-Forward Network*.

Packets from each route r arrive, respectively, as a Poisson processes of rate a_r and join the first queue on route r , namely, j_1^r . Packets at each queue are assumed to have a

service requirement that is independent exponentially distributed of mean 1. Given \mathbf{Q} packets are served from queue j at rate $\sigma_j^{SF}(\mathbf{Q})$. The jobs within a queue are served in a first-in first-out manner and a job completing service at the front of the queue will subsequently join the back of the next queue on its route (or depart the network if the packet is at the final queue on its route).

B. Kelly Networks, Store-Forward and $\Phi(\mathbf{Q})$

We now give an explicit expression for the Store-Forward allocation and the associated function Φ . We do so by explaining a seemingly unrelated model of a closed queueing network.

Consider a network of processor-sharing queues indexed by \mathcal{L} . Each server processes work at a unit rate. To avoid ambiguity with the index j , we henceforth refer to these queues $l \in \mathcal{L}$ as *pools*. We assume that jobs from classes indexed by \mathcal{J} arrive into this network at rate a_j , $j \in \mathcal{J}$. We assume that the pools are ordered, $\mathcal{L} = \{l_1, l_2, \dots, l_{|\mathcal{L}|}\}$, and that each arriving job visits these queues sequentially in this order. The load of a job at a queue depends on the class j and the index l of the pool. In particular, we assume that a class j job at queue l has a service requirement that is exponentially distributed with mean A_{lj}^{-1} .

It is somewhat classical queueing theory that the stationary number of jobs of each class at each pool, $\mathbf{m} = (m_{lj} : j \in \mathcal{J}, l \in \mathcal{L})$, is

$$\pi^\circ(\mathbf{m}) = \prod_{l \in \mathcal{L}} \left[(1 - a_l)^{-1} \binom{m_l}{m_{lj} : j \in \mathcal{J}} \prod_{j \in \mathcal{J}} (A_{lj} a_j)^{m_{lj}} \right]. \quad (3)$$

We refer the reader to Kelly [9, Chapter 3] for a standard treatment of this result. We condition the number of packets of each class to be $\mathbf{Q} = (Q_j : j \in \mathcal{J})$ then the (conditional) stationary distribution of packets in this network is given by

$$\pi^\circ(\mathbf{m}|\mathbf{Q}) = \frac{1}{\Phi(\mathbf{Q})} \prod_{l \in \mathcal{L}} \left[\binom{m_l}{m_{lj} : j \in \mathcal{J}} \prod_{j \in \mathcal{J}} (A_{lj})^{m_{lj}} \right].$$

where $\Phi(\mathbf{Q})$, as was briefly introduced in the previous subsection, is defined to be the normalizing constant achieved by summing the above terms over \mathbf{m} with $\sum_{l:j \in \mathcal{J}} m_{lj} = Q_j$, $j \in \mathcal{J}$ (after canceling out the a_j terms). Namely,

$$\Phi(\mathbf{Q}) \prod_{j \in \mathcal{J}} (a_j^{Q_j}) = \pi^\circ\left(\left\{\mathbf{m} : \sum_{l:j \in \mathcal{J}} m_{lj} = Q_j, j \in \mathcal{J}\right\}\right). \quad (4)$$

We now have a formal definition for $\Phi(\mathbf{Q})$; however the significance of the constant is its relationship with the throughput with the network just described. Basically, conditional the number of customers of each class $\mathbf{Q} = (Q_j : j \in \mathcal{J})$, the stationary throughput for class j jobs in this processor-sharing network is given by the ratio

$$\frac{\Phi(\mathbf{Q} - \mathbf{e}_j)}{\Phi(\mathbf{Q})}, \quad (5)$$

in other words, by the Store-Forward allocation, (2). It is well known that allocations of the above form, admit reversible and insensitive Markov chains, see Whittle [25].

For these calculations, which are only briefly described, are relatively straight-forward and we refer the reader to [9], [23] for detailed proofs.

C. Proportional Fairness and Store-Forward

Given a vector of queue sizes $\mathbf{Q} = (Q_j : j \in \mathcal{J})$, the proportional fair optimization is

$$\sigma^{PF}(\mathbf{Q}) \in \max_{\sigma \in \langle \mathcal{S} \rangle} \sum_{j \in \mathcal{J}} Q_j \log \sigma_j. \quad (6)$$

The proportional fair optimization was first described by Kelly [11]. From this optimization, we can define the Proportional Scheduler for multi-hop switch networks. We do this in the next subsection. The main connection between Store-Forward and proportional fairness is that the limit of the Store-Forward allocation is a solution to Proportional fair optimization.

Proposition 1:

$$\sigma^{SF}(c\mathbf{Q}) \xrightarrow{c \rightarrow \infty} \sigma^{PF}(\mathbf{Q})$$

This result is proven in [23] and earlier heuristic derivations can be found in Schweitzer [16], Kelly [10], Roberts and Massoulié [13]. These statements generalize to further reversible allocations Massoulié [12] and Walton [22].

D. BackPressure and the Proportional Scheduler

For multi-hop networks, the BackPressure policies represent the canonical maximum stable scheduling policy. Letting $X_{jr}(t)$ be the number of route r packets at queue j at time t , the BackPressure policy is as follows

BP1. Calculate weights by comparing with downstream queue lengths,

$$w_j(X(t)) = \max_{r:j \in r} \left\{ X_{jr}(t) - X_{j_+^r}(t), 0 \right\},^2 \quad (7)$$

and let $r_j^{BP}(t+1)$ be the solution to this maximization. Here j_+^r denotes the queue subsequent to j on route r .

BP2. Over set of schedules \mathcal{S} , solve the optimization

$$\max_{\sigma \in \mathcal{S}} \sum_{j \in \mathcal{J}} \sigma_j w_j(X(t)), \quad (8)$$

and let $\sigma^{BP}(t+1)$ be an optimal solution.

BP3. If $w_j(X(t)) > 0$, at the next time instance schedule $\sigma_j^{BP}(t+1)$ packets from route $r_j^{BP}(t+1)$ from each queue $j \in \mathcal{J}$, else, do not schedule any packets from queue j .

Notice in the first step above, information must be exchanged along links to make a queue size comparison (7). Further, note that the policy needs to know the route of each packet at each queue in order to make a scheduling decision. If the number of routes is large, then this can be prohibitive.

²Here if there is no next link after j on route r then we set $X_{j_+^r} = 0$.

However, it is argued in [24], [2] that there are substantial advantages in implementing a proportional fair optimization in a switch network. For the multi-hop switch networks with fixed routing, the Proportional Scheduler is as follows: given a vector of link queue sizes $(Q_j(t) : j \in \mathcal{J})$,

PS1. Over set of schedules $\langle \mathcal{S} \rangle$ solve the optimization

$$\text{maximize} \quad \sum_{j \in \mathcal{J}} Q_j(t) \log(s_j) \quad (9a)$$

$$\text{over} \quad s \in \langle \mathcal{S} \rangle. \quad (9b)$$

Let $\sigma(t+1)$ be a random variable on \mathcal{S} whose mean solves this optimization.

PS2. From each queue $j \in \mathcal{J}$, serve $\sigma_j(t+1)$ packets from the front of queue j . Each of these packets then join the back of their next downstream queue as determine by its route class.³

It is proven in the recent paper, [2], that for FIFO service the proportional scheduler is maximum stable. When the rule PS2 is replaced with a random service discipline, maximum stability is proven in the paper [24]. The main structural advantage of the Proportional Scheduler over BackPressure is that it does not require information about the route taken by a packet in order to make a scheduling decision. This substantially reduces the information required for implementation and thus exhibits great potential for decentralized implementations. See [24] and [2] for a more in depth discussion of this point.

E. Switch networks and Store-Forward

Switch networks as described in Section III-D are discrete time queueing networks while the Store-Forward Network as defined in Section III-A is a continuous time queue network. Thus, although there is an asymptotic relationship between the Proportional Scheduler and Store-Forward, the Store-Forward allocation is not constructed to be an implementable discrete time scheduling policy for switch networks and is instead used for reasons of analytic tractability. Nonetheless, we note that implementations of Store-Forward are possible in switch networks, we refer the read to [18] for an algorithm which implements a continuous time Store-Forward network on a single-hop switched queueing network. Here certain optimality properties can be proven due to this relationship with product form queue networks.

IV. CALCULATIONS FOR STORE-FORWARD NETWORKS

In this section we derived various stationary characteristics for FIFO routed Store-Forward Networks: in Section IV-A, we give the stationary distribution of this Store-Forward Network; in Section IV-B we give the stationary route delay of a Store-Forward Network; in Section IV-C we derive product form results for the stationary Store-Forward network and analyse it on independent set scheduling for perfect graphs; and finally, in Section IV-D we heuristically a large deviations principle for these networks and discuss

how it can be used to construct a Lyapunov function used for stability analysis. For each result, we conjecture an analogous result for the Proportional Scheduler.

A. Stationary Distribution and Stability

It is not hard to verify that the Store-Forward Network as described in Section III-A is quasi-reversible and so by application of Kelly's lemma [9, Theorem 1.13] has stationary distribution as follows:

Theorem 1: A FIFO routed Store-Forward Network is positive recurrent when $(a_j : j \in \mathcal{J}) \in \langle \mathcal{S} \rangle^\circ$ and has an equilibrium distribution of the form

$$\pi(\mathbf{Q}, \mathbf{\Gamma}) = \Phi(\mathbf{Q}) \prod_{j \in \mathcal{J}} \prod_{r: j \in r} \left(a_r^{\Gamma_{jr}(Q_j)} \right). \quad (10)$$

The proof of this Theorem is given in the appendix. Note that an immediate consequence of the above theorem is the following:

Corollary 1: For a stationary FIFO routed Store-Forward Network, conditional on length of queue j , the route-class of a packet at the queue is independent and from route r with probability

$$\frac{a_r}{a_j}, \quad r \ni j.$$

An important consequence of the structure of the Store-Forward allocation is that the stationary queue size \mathbf{Q} can be represented in terms of a number of independent random variables.

Proposition 2: There are mutually independent random vectors $\mathbf{m}_l = (m_{lj} : j \in l)$, $l \in \mathcal{L}$, where

$$Q_j = \sum_{l: l \ni j} m_{lj},$$

and each

$$m_l := \sum_{j \in l} m_{lj}, \quad l \in \mathcal{L},$$

is a geometric random variable with parameter a_l and, conditional on m_l , the random variables $(m_{lj} : j \in l)$ have a multinomial distribution with m_l trials and parameters

$$\left(\frac{A_{lj} a_j}{a_l} : j \in l \right).$$

A proof of this result is given in the Appendix.

The above proposition can be observed directly from the definition of the Store-Forward allocation. Its consequences for proportional fair systems are first explored by Kang, Kelly, Lee and Williams [7] in order to analyse resource pooling effects in Bandwidth sharing networks with multi-path routing. It is first applied to switch systems by Shah et al [18] to prove optimal scaling behaviour for switch networks in heavy traffic.

B. Delay

With Theorem 1 and Proposition 2, we can analyse the delay of a packet traversing its route in a Store-Forward network as follows.

³In joining the back of queue, ties are broken arbitrarily.

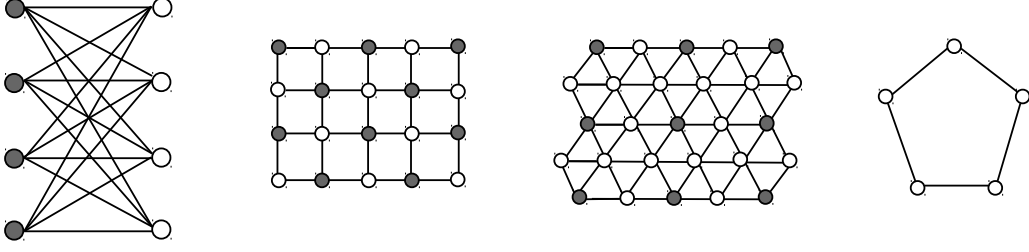


Fig. 1. Four examples of interference graphs: a complete bipartite graph (input-queue switch), a square grid, a triangular grid and a odd length cycle. The first three networks are perfect graphs. In each example, the nodes coloured grey are mutually independent.

Theorem 2: For a stationary Store-Forward network, the delay on a route r given by D_r has expectation

$$\mathbb{E}[D_r] = \sum_{j \in r} \sum_{l: j \in l} \frac{A_{lj}}{1 - a_l},$$

or, as a shorthand in matrix multiplication, letting $\bar{\mathbf{m}} = ((1 - a_l)^{-1} : l \in \mathcal{L})$ and $\bar{\mathbf{n}}^r = (\mathbb{I}[j \in r] : j \in \mathcal{J})$ then

$$\mathbb{E}[D_r] = \bar{\mathbf{m}}^\top \mathbf{A} \bar{\mathbf{n}}^r.$$

Proof: First by Little's Law the delay on a route, D_r , has expectation

$$\mathbb{E}[D_r] = \frac{1}{a_r} \sum_{j \in r} \mathbb{E}[Q_{jr}].$$

By Corollary 1 above,

$$\mathbb{E}[Q_{jr}] = \frac{a_r}{a_j} \mathbb{E}[Q_j]$$

and by Proposition 2, we know that

$$\mathbb{E}[Q_j] = \sum_{l: j \in l} \mathbb{E}[m_{lj}] = \sum_{l: j \in l} \frac{A_{lj} a_j}{a_l} \mathbb{E}[m_l]$$

and

$$\mathbb{E}[m_l] = \frac{a_l}{1 - a_l}.$$

Combining the four equalities above, we have as required,

$$\mathbb{E}[D_r] = \sum_{j \in r} \sum_{l: j \in l} \frac{A_{lj}}{1 - a_l}.$$

Remark 1: The result generalizes if we allow routes to visit queues more than once, in particular, if we let $\bar{\mathbf{m}} = ((1 - a_l)^{-1} : l \in \mathcal{L})$ and we let $\bar{\mathbf{n}}^r = (n_j^r : j \in \mathcal{J})$ give the (mean) number of times a route r job visits queue j then the delay on that route is given by matrix multiplication

$$\mathbb{E}[D_r] = \bar{\mathbf{m}}^\top \mathbf{A} \bar{\mathbf{n}}^r.$$

Remark 2: It is well known that the delay for the BackPressure policy depends on length of the route, see [4], [19] Note that unlike BackPressure, the position of a queue within a route does not effect the queue size and delay, what matters is the delay induced by the load vector $(a_l : l \in \mathcal{L})$.

This is significant because if one makes an improvement in the service capacity in parts of the network then the queue sizes and, thus, delay will be reduced in that part of the network. For example, consider a single long route where schedules between queues do not interfere. For BackPressure under moderate load it is known that queue sizes grow linearly from the last queue to the first [4], and an improvement in the service capacity of a queue – say we double the service capacity of the first queue on the route – will not reduce queue sizes below the size of the queue in front. However, the Store-Forward network we see from the above expression that doubling service capacity will halve the queue size at that queue and thus significantly reduce delays.

Remark 3: Given Theorem 2, it is reasonable to conjecture that the same holds for the Proportional Scheduler, that is, if D_r^{PS} gives the stationary route delay of the proportional scheduler, then

$$\mathbb{E}[D_r^{PS}] \sim \sum_{j \in r} \sum_{l: j \in l} \frac{A_{lj}}{1 - a_l} \quad (11)$$

as we let the load vector $(a_j : j \in \mathcal{J})$ approach the boundary of the scheduling polytope $\langle \mathcal{S} \rangle$.

To understand the effect of loads on queue sizes a better understanding of the effect the scheduling set \mathcal{S} on the matrix \mathbf{A} is required. For certain examples, bipartite graphs and perfect graphs, the matrix \mathbf{A} has a relatively simple structure we discuss this in the next section.

C. Product Form Resource Pooling

The underlying queueing mechanism for Store-Forward is quasi-reversible. It is often found that quasi-reversible Markov chains exhibit product-form stationary distributions on product sets. This is observed to be the case for Store-Forward. For instance, in the upcoming paper [2, Proposition 1] it is proven that a Store-Forward network whose schedules are a product set, i.e., $\mathcal{S} = \mathcal{S}_1 \times \dots \times \mathcal{S}_N$, the stationary queue size vectors associated with each of these components are independent.

We generalize this result in the following proposition. Typically product-form results apply over product sets; however, an interesting consequence of this result is that the

queues considered do not need the scheduling set \mathcal{S} to be of a product type in order to have product-form stationary behaviour.

Theorem 3: Consider a stationary Store-Forward Network on scheduling set

$$\langle \mathcal{S} \rangle = \left\{ \mathbf{s} \in \mathbb{R}_+^{\mathcal{J}} : \sum_{j \in \mathcal{J}} A_{lj} s_j \leq 1, l \in \mathcal{L} \right\}. \quad (12)$$

If there are two queues j and j' such that there is no share resource pool, i.e. $\nexists l \in \mathcal{L}$ such that $A_{lj} > 0$ and $A_{lj'} > 0$, then the queues are statistically independent.

Note when we say the queues are independent we also mean the route classes of the packets within the queue are also statistically independent. A proof of this result is given in the Appendix. Further, the result immediately extends to give the independence of sets of queues provided no pair of queues shares a common resource pool.

Now the above result may seem somewhat abstract because of the dependency on the matrix \mathbf{A} . A good example to consider is where the matrix \mathbf{A} corresponds to the interference graph of a perfect graph. We take a graph $G = (\mathcal{J}, \mathcal{E})$ with vertices \mathcal{J} and edges \mathcal{E} . Here a queue located at each vertex of the graph can transmit a packet provided none of its neighbours transmit. Thus the set of schedules are the independent sets of this graph:

$$\mathcal{S} = \{ \sigma \in \mathbb{Z}_+^{\mathcal{J}} : \sigma_j + \sigma_{j'} \leq 1, (j, j') \in \mathcal{E} \}$$

A graph is perfect if the neither the graph G nor its complement contain an odd cycle of length 5 or greater. Here bipartite graphs are an important special case. When a graph is perfect the convex hull of \mathcal{S} takes a explicit form

$$\langle \mathcal{S} \rangle = \left\{ \mathbf{s} \in [0, 1]^{\mathcal{J}} : \sum_{j \in \mathcal{C}} s_j \leq 1, \mathcal{C} \in \mathcal{C} \right\}$$

where \mathcal{C} gives the set of cliques of the graph G . This deep result is proven in [5]. So for a perfect graph, the resource pools are the cliques of the graph and to queues do not share a clique so long as they are not neighbours. Thus a direct consequence of the above theorem is the following observation

Corollary 2: For a Store-Forward network on an interference graph which is perfect, if two queues are not neighbours then they are statistically independent.

See Figure 1 for some example of perfect graphs. The result is interesting since these graphs appear to exhibit periods of independence when mixing between disjoint modes of operation under different scheduling algorithms, e.g. for a complete bipartite graph either the left-hand side is sending or the right-side and during these periods queue sizes at nodes should be approximately independent. See [27] and [28] for in depth discussion a CSMA scheduling algorithm on these network topologies.

Of course, it is natural to conjecture that this same independent behaviour occurs for the Proportional Scheduler under limit regimes where the network is congested, such as heavy traffic or large deviations limits.

D. Large Deviations Estimates and Lyapunov functions

Finally, we analyse the large deviations behaviour of the Proportional Scheduler. Analogous to Massoulié [12], the large deviations rate function found in Theorem 1 provides a Lyapunov function that we can then use to prove stability for the Proportional Scheduler.⁴ Indeed, the hope is that calculations of this type can be used to form Lyapunov functions of other FIFO switch scheduling policies. However, currently it is not clear how critical the underlying reversibility of the model is required in order for this analysis to work out.

For reasons of space, we give a heuristic derivation of the rate function associated with the stationary distribution, (10). We take a sequence of states $(\mathbf{Q}^c, \mathbf{\Gamma}^c)$ where

$$\frac{\mathbf{Q}^c}{c} \xrightarrow{c \rightarrow \infty} \mathbf{Q}$$

and for a piecewise linear process $\mathbf{\Gamma}$

$$\frac{\mathbf{\Gamma}^c(cq)}{c} \xrightarrow{c \rightarrow \infty} \mathbf{\Gamma}(q),$$

for $q \leq \mathbf{Q}_j$. Here, convergence is point-wise (or uniform since the limit process is increasing and continuous). Since $\mathbf{\Gamma}$ is assumed to be piecewise linear, we let $k = 0, \dots, K$ index these linear stages with each stage each started at queue sizes $Q_j(k)$, $k = 0, \dots, K$. Let $\Delta Q_j(k) = (Q_j(k) - Q_j(k-1))$ and let $\mathbf{\Gamma}'_{jr}(k)$ index the gradient of $\mathbf{\Gamma}$ in each stage. By definition

$$\sum_r \mathbf{\Gamma}'_{jr}(k) = 0$$

for each $j \in \mathcal{J}$ and $k = 1, \dots, K$. So we may interpret these gradients as probability distributions giving the relative density of packets along each FIFO queue.

Under π , the stationary probability that $\mathbf{\Gamma}^c(Q_j^c(k))/c \approx \mathbf{\Gamma}(Q_j(k))$, for $k = 0, \dots, K$ and that $\mathbf{Q}^c \approx c\mathbf{Q}$ is approximated by

$$P^c := \Phi(\mathbf{Q}^c) \prod_{j \in \mathcal{J}} \prod_{k=1}^K \left(\frac{c \Delta Q_j(k)}{c \mathbf{\Gamma}'_{jr} \Delta Q_j(k) : r \ni j} \right) \prod_{r \in \mathcal{R}} a_r^{\mathbf{\Gamma}'_{jr} \Delta Q_j(k)}.$$

It can be shown that $\Phi(\mathbf{Q}^c)$ approximates the solution of the proportional fair optimization in the following sense

$$\lim_{c \rightarrow \infty} \frac{1}{c} \log \Phi(\mathbf{Q}^c) = - \max_{\sigma \in \langle \mathcal{S} \rangle} \sum_{j \in \mathcal{J}} Q_j \log \sigma_j.$$

A short argument for the above equality can be found in [22]. Applying a Stirling's approximation to the multinomial

⁴Theorem 1 was proven in order to justify proceeding with the proofs in [24] and [2].

term about yields the expression

$$\begin{aligned}
& \lim_{c \rightarrow \infty} \frac{1}{c} \log P^c \\
&= - \max_{\sigma \in \langle S \rangle} \sum_{j \in \mathcal{J}} Q_j \log \sigma_j \\
&+ \sum_j \sum_k \left[\Delta Q_j(k) \log \Delta Q_j(k) \right. \\
&\quad \left. - \sum_r \Gamma'_{jr} \Delta Q_j(k) \log \left(\Gamma'_{jr} \Delta Q_j(k) \right) \right. \\
&\quad \left. + \sum_r \Gamma'_{jr} \Delta Q_j(k) \log a_r \right] \\
&= - \max_{\sigma \in \langle S \rangle} \sum_{j \in \mathcal{J}} Q_j \log \sigma_j \\
&\quad - \sum_j \sum_r \sum_k \Delta Q_j(k) \Gamma'_{jr}(k) \log \left(\frac{\Gamma'_{jr}(k)}{a_r} \right) \\
&= - \max_{\sigma \in \langle S \rangle} \sum_{j \in \mathcal{J}} Q_j \log \sigma_j - \sum_j \sum_r \int_0^{Q_j} \log \left(\frac{\Gamma'_{jr}}{a_r} \right) d\Gamma_{jr}
\end{aligned}$$

The entropy terms derived above can be seen as a form of Sanov's Theorem. The rate function with this integral representation is used by Bramson as a Lyapunov function to prove stability for FIFO queueing networks with fixed service capacity [3]. Removing this integral term, a similar large deviations argument is used by Massoulié to prove stability for proportional fair networks with probabilistic routing [12]. Combining these entropy arguments, the above rate function provides a Lyapunov function for the Proportional Scheduler in FIFO switch networks, [2]. It is natural to conjecture that the above rate function is the large deviations rate function for a FIFO network operating under the Proportional Scheduler. It is currently unclear the extent to which the above heuristic and Lyapunov function can be developed beyond a proportional fair framework.

REFERENCES

- [1] T. Bonald and A. Proutière, "On performance bounds for balanced fairness," *Performance Evaluation*, vol. 55, pp. 25–50, 2004.
- [2] M. Bramson, B. D'Auria, and N. Walton, "Proportional switching in fifo routed networks," *in preparation*.
- [3] M. Bramson, "Convergence to equilibria for fluid models of FIFO queueing networks," *Queueing Systems Theory Appl.*, vol. 22, no. 1-2, pp. 5–45, 1996.
- [4] L. Bui, R. Srikant, and A. L. Stolyar, "A novel architecture for reduction of delay and queueing structure complexity in the back-pressure algorithm," *IEEE/ACM Trans. Netw.*, vol. 19, no. 6, pp. 1597–1609, 2011.
- [5] M. Chudnovsky, N. Robertson, P. Seymour, and R. Thomas, "The strong perfect graph theorem," *Annals of Mathematics*, pp. 51–229, 2006.
- [6] B. Ji, C. Joo, and N. Shroff, "Throughput-optimal scheduling in multi-hop wireless networks without per-flow information," *Networking, IEEE/ACM Transactions on*, vol. 21, no. 2, pp. 634–647, April 2013.
- [7] W. N. Kang, F. P. Kelly, N. H. Lee, and R. J. Williams, "Product form stationary distributions for diffusion approximations to a flow level model operating under a proportional fair sharing policy," *Performance Evaluation Review*, vol. 35, pp. 36–38, 2007.
- [8] —, "State space collapse and diffusion approximation for a network operating under a fair bandwidth sharing policy," *Ann. Appl. Probab.*, vol. 19, pp. 1719–1780, 2009.

- [9] F. P. Kelly, *Reversibility and Stochastic Networks*. Chichester: Wiley, 1979.
- [10] —, "On a class of approximations for closed queueing networks," *Queueing Systems*, vol. 4, pp. 69–76, 1989.
- [11] —, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, pp. 33–37, 1997.
- [12] L. Massoulié, "Structural properties of proportional fairness: Stability and insensitivity," *Ann. Appl. Probab.*, vol. 17, pp. 809–839, 2007.
- [13] L. Massoulié and J. Roberts, "Bandwidth sharing: Objectives and algorithms," *IEEE Infocom 1999*, vol. 10, no. 3, pp. 320–328, 1999.
- [14] N. McKeown, A. Mekkittikul, V. Anantharam, and J. Walrand, "Achieving 100% throughput in an input-queued switch," *Communications, IEEE Transactions on*, vol. 47, no. 8, pp. 1260–1267, 1999.
- [15] M. Neely, E. Modiano, and C. Rohrs, "Dynamic power allocation and routing for time-varying wireless networks," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 1, pp. 89–103, Jan 2005.
- [16] P. J. Schweitzer, "Approximate analysis of multiclass closed networks of queues," *Proceedings of the international conference on stochastic control and optimization*, 1979.
- [17] D. Shah, J. Tsitsiklis, and Y. Zhong, "Optimal scaling of average queue sizes in an input-queued switch: an open problem," *Queueing Systems*, pp. 1–10, 2011.
- [18] D. Shah, N. S. Walton, and Y. Zhong, "Optimal queue-size scaling in switched networks," *The Annals of Applied Probability*, vol. 24, no. 6, pp. 2207–2245, 12 2014.
- [19] A. L. Stolyar, "Large number of queues in tandem: Scaling properties under back-pressure algorithm," *Queueing Syst. Theory Appl.*, vol. 67, no. 2, pp. 111–126, Feb. 2011.
- [20] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *Automatic Control, IEEE Transactions on*, vol. 37, no. 12, pp. 1936–1948, 1992.
- [21] M. Vlasios, J. Zhang, and B. Zwart, "Insensitivity of proportional fairness in critically loaded bandwidth sharing networks," (*preprint*), 2014.
- [22] N. Walton, "Insensitive, maximum stable allocations converge to proportional fairness," *Queueing Systems*, vol. 68, pp. 51–60, 2011.
- [23] N. S. Walton, "Proportional fairness and its relationship with multi-class queueing networks," *Ann. Appl. Probab.*, vol. 22, no. 6, pp. 2301–2333, 2009.
- [24] —, "Concave switching in single and multihop networks," in *The 2014 ACM International Conference on Measurement and Modeling of Computer Systems*, ser. SIGMETRICS '14. New York, NY, USA: ACM, 2014, pp. 139–151.
- [25] P. Whittle, "Partial balance and insensitivity," *J. Appl. Probab.*, vol. 22, pp. 168–176, 1985.
- [26] H.-Q. Ye and D. D. Yao, "A stochastic network under proportional fair resource control diffusion limit with multiple bottlenecks," *Operations Research*, vol. 60, no. 3, pp. 716–738, 2012.
- [27] A. Zocca, S. Borst, J. van Leeuwen, and F. Nardi, "Delay performance in random-access grid networks," *Performance Evaluation*, vol. 70, no. 10, pp. 900 – 915, 2013, proceedings of {IFIP} Performance 2013 Conference.
- [28] A. Zocca, S. C. Borst, and J. S. van Leeuwen, "Mixing properties of csma networks on partite graphs," in *Performance Evaluation Methodologies and Tools (VALUETOOLS), 2012 6th International Conference on*. IEEE, 2012, pp. 117–126.

APPENDIX

We provide proofs of a number of results in the body of the text. First, we prove Theorem 1.

Proof: [Proof of Theorem 1] We verify that $\pi(\mathbf{Q}, \Gamma)$ is a stationary distribution by confirming quasi-reversibility of our Markov chain. Here we must define the time-reversal of the Store-Forward network. In this time reversal, packets on each route r arrive as a Poisson process of rate a_r at the last queue on route $r = (j_1^r, \dots, j_{k_r}^r)$, namely $j_{k_r}^r$. Packets have a service requirement that is independent exponentially distributed with mean 1. Given the vector of queue sizes \mathbf{Q}

packets, queue are served at a rate as given by the Store-Forward allocation, $\sigma^{SF}(\mathbf{Q})$. Queues are served in a FIFO order; however, in comparison to the Store-Forward network, (in forward time) jobs are served from the end of the queue and arrival are placed at the front of the queue. A packet of route r completing service on the k th stage of its route, i.e. at queue j_k^r , then moves to the $k-1$ th queue on its route, i.e. to queue j_{k-1}^r , or leaves the network if $k=1$ and thus has completed service at all queues on its route.

Now there are three types of transition that can occur: an arrival on route r ; a departure on route r ; and a transition between of a route r job between queues j and j' . We let $q((\mathbf{Q}, \Gamma), (\mathbf{Q}', \Gamma'))$ give the transition rates of the Store-Forward network as described in Section III-A. We let $q^R((\mathbf{Q}, \Gamma), (\mathbf{Q}', \Gamma'))$ be the transition rates of the time reversal of this Store-Forward network, as described in the above paragraph.

We verify balance equations, first, for an arrival transition on route r at queue $j = j_1^r$, here a transition from (\mathbf{Q}, Γ) to (\mathbf{Q}', Γ') where \mathbf{Q} occurs where $Q_j' = Q_j + 1$ and $\Gamma_{j,r}'(Q_j + 1) = \Gamma_{j,r}(Q_j) + 1$, all other components of (\mathbf{Q}, Γ) and (\mathbf{Q}', Γ') are equal.

$$\pi(\mathbf{Q}, \Gamma) \times q((\mathbf{Q}, \Gamma), (\mathbf{Q}', \Gamma')) \quad (13)$$

$$= \underbrace{\frac{\Phi(\mathbf{Q})}{\Phi(\mathbf{Q} + \mathbf{e}_j)}}_{\sigma_j^{SF}(\mathbf{Q}')} \Phi(\mathbf{Q} + \mathbf{e}_j) \prod_{j \in \mathcal{J}} \prod_{r: j \in r} \left(a_r^{\Gamma_{j,r}(Q_j)} \right) \times a_r \quad (14)$$

$$= \pi(\mathbf{Q}', \Gamma') \times q^R((\mathbf{Q}', \Gamma'), (\mathbf{Q}, \Gamma)) \quad (15)$$

The same argument taken from equation (15) to (13) gives the equivalent expression shows that these balance equations hold for a departure transition. Next, for a transition between queues j and j' on route r , a transition from (\mathbf{Q}, Γ) to (\mathbf{Q}', Γ') occurs with $Q_{j'}' = Q_{j'} + 1$, $Q_j' = Q_j - 1$, $\Gamma_{j',r}'(Q_{j'} + 1) = \Gamma_{j',r}(Q_{j'}) + 1$ and $\Gamma_{j,r}'(Q_j) = \emptyset$, all other components of (\mathbf{Q}, Γ) and (\mathbf{Q}', Γ') are equal.

$$\pi(\mathbf{Q}, \Gamma) \times q((\mathbf{Q}, \Gamma), (\mathbf{Q}', \Gamma')) \quad (16)$$

$$= \Phi(\mathbf{Q}) \prod_{j \in \mathcal{J}} \prod_{r: j \in r} \left(a_r^{\Gamma_{j,r}(Q_j)} \right) \times \frac{\Phi(\mathbf{Q} - \mathbf{e}_j)}{\Phi(\mathbf{Q})} \quad (17)$$

$$= \Phi(\mathbf{Q}') \prod_{j \in \mathcal{J}} \prod_{r: j \in r} \left(a_r^{\Gamma_{j,r}'(Q_{j'}')} \right) \times \frac{\Phi(\mathbf{Q}')}{\Phi(\mathbf{Q}' + \mathbf{e}_j)} \quad (18)$$

$$= \pi(\mathbf{Q}', \Gamma') \times q^R((\mathbf{Q}', \Gamma'), (\mathbf{Q}, \Gamma)) \quad (19)$$

Finally, for each \mathbf{Q} ,

$$\begin{aligned} \sum_{\mathbf{Q}', \Gamma'} q((\mathbf{Q}, \Gamma), (\mathbf{Q}', \Gamma')) &= \sum_r a_r + \sum_j \sigma_j^{SF}(\mathbf{Q}) \\ &= \sum_{\mathbf{Q}', \Gamma'} q^R((\mathbf{Q}, \Gamma), (\mathbf{Q}', \Gamma')). \end{aligned}$$

This verifies that the conditions of Kelly's Lemma [9, Theorem 3.1] hold and thus (10) gives the stationary distribution of our FIFO routed Store-Forward network.

Finally we show that the measure π can be normalized,

$$\begin{aligned} \sum_{\mathbf{Q}, \Gamma} \pi(\mathbf{Q}, \Gamma) &= \sum_{\mathbf{Q}} \Phi(\mathbf{Q}) \prod_{j \in \mathcal{J}} a_j^{Q_j} \\ &= \sum_{\mathbf{Q}} \pi^\circ \left(\left\{ \mathbf{m} : \sum_{l: j \in l} m_{jl} = Q_j, j \in \mathcal{J} \right\} \right) = 1 \end{aligned}$$

where in the third inequality we recall (4). \blacksquare

We now provide a proof of Proposition 2 which established a product form relation in Store-Forward Networks.

Proof: [Proof of Proposition 2] It is clear that for the distribution $\pi^\circ(\mathbf{m})$, (3), gives the distribution of the random variables \mathbf{M}_l as described in Proposition 2. And, as given in (4), for this distribution the following identity holds

$$\Phi(\mathbf{Q}) \prod_{j \in \mathcal{J}} \left(a_j^{Q_j} \right) = \pi^\circ \left(\left\{ \mathbf{m} : \sum_{l: j \in l} m_{jl} = Q_j, j \in \mathcal{J} \right\} \right). \quad (20)$$

Further, the following identity holds for $\pi(\mathbf{Q})$, the stationary distribution of a Store-Forward network (10) ignoring the effect of the distribution of packets in the queue, Γ ,

$$\begin{aligned} \pi(\mathbf{Q}) &:= \sum_{\Gamma} \pi(\mathbf{Q}, \Gamma) \\ &= \sum_{\Gamma} \Phi(\mathbf{Q}) \prod_{j \in \mathcal{J}} \prod_{k=1}^{Q_j} \prod_{r \in \mathcal{R}} a_r^{\Gamma_{j,r}(k) - \Gamma_{j,r}(k-1)} \\ &= \Phi(\mathbf{Q}) \prod_{j \in \mathcal{J}} \prod_{k=1}^{Q_j} \left(\sum_{r: j \in r} a_r \right) \\ &= \pi^\circ \left(\left\{ \mathbf{m} : \sum_{l: j \in l} m_{jl} = Q_j, j \in \mathcal{J} \right\} \right). \end{aligned}$$

where the final equality about uses identity (20). This established the relationship between \mathbf{Q} and independent random variables \mathbf{m}_l , $l \in \mathcal{L}$, as described. \blacksquare

Proof: [Proof of Theorem 3] From Proposition 2, we know that the stationary queue sizes of the Store-Forward Network, \mathbf{Q} , relate to the stationary distribution of the product form queue network, \mathbf{m} , through the equality

$$Q_j = \sum_{l: l \in j} m_{lj}.$$

We recall the notation that $l \in j$ when $A_{lj} > 0$ and $l \notin j$ if $A_{lj} = 0$. Since the random vectors $\mathbf{m}_l = (m_{lj} : j \in \mathcal{J})$ are mutually independent over $l \in \mathcal{L}$. We observe that if j and j' do not share a common resource pool, then the summations

$$Q_j = \sum_{l: l \in j} m_{lj}, \quad Q_{j'} = \sum_{l: l \in j'} m_{lj'}$$

both sum over mutually exclusive indices l . Thus the queue sizes Q_j and $Q_{j'}$ are independent. Further, since the distribution of jobs within each queue $\Gamma_{j,r}$ and $\Gamma_{j',r}$ are independent when we condition on Q_j and $Q_{j'}$, the queue size distribution of the queues are independent. \blacksquare